

PHNL031092

PCT/IB2004/051560

1

Robust de-interlacing of video signals

The invention relates to a method for de-interlacing, in particular GST-based de-interlacing a video signal with estimating a motion vector for pixels from said video signal, defining a current field of input pixels from said video signal to be used for calculating an interpolated output pixel, and calculating an interpolated output pixel from a weighted sum of said input pixels. The invention further relates to a display device and a computer program for de-interlacing a video signal.

De-interlacing is the primary resolution determination of high-end video display systems to which important emerging non-linear scaling techniques such as DRC and Pixel Plus, can only add finer detail. With the advent of new technologies like LCD and PDP, the limitation in the image resolution is no longer in the display device itself, but rather in the source or transmission system. At the same time these displays require a progressively scanned video input. Therefore, high quality de-interlacing is an important pre-requisite for superior image quality in such display devices.

15

A first step to de-interlacing is known from P. Delonge, et al., "Improved Interpolation, Motion Estimation and Compensation for Interlaced Pictures", IEEE Tr. on Im. Proc., Vol. 3, no. 5, Sep. 1994, pp 482-491.

The disclosed method is also known as the general sampling theorem (GST) de-interlacing method. The method is depicted in Fig. 1. Fig. 1 depicts a field of pixels 2 in a vertical line on even vertical positions $y+4 - y-4$ in a temporal succession of $n-1 - n$. For de-interlacing, two independent sets of pixel samples are required. The first set of independent pixel samples is created by shifting the pixels 2 from the previous field $n-1$ over a motion vector 4 towards a current temporal instance n into motion compensated pixel samples 6. The second set of pixels 8 is also located on odd vertical lines $y+3 - y-3$. Unless the motion vector 6 is small enough, e.g. unless a so-called "critical velocity" occurs, i.e. a velocity leading to an odd integer pixel displacements between two successive fields of pixels, the pixel samples 6 and the pixels 8 are assumed to be independent. By weighting the

pixel samples 6 and the pixels 8 from the current field the output pixel sample 10 results as a weighted sum (GST-filter) of samples.

Mathematically, the output sample pixel 10 can be described as follows. Using $F(\bar{x}, n)$ for the luminance value of a pixel at position \bar{x} in image number n , and using F_i for the luminance value of interpolated pixels at the missing line (e.g. the odd line) the output of the GST de-interlacing method is as:

$$F_i(\bar{x}, n) = \sum_k F(\bar{x} - (2k+1)\bar{u}_y, n)h_1(k, \delta_y) + \sum_m F(\bar{x} - \bar{e}(\bar{x}, n) - 2m\bar{u}_y, n-1)h_2(m, \delta_y)$$

with h_1 and h_2 defining the GST-filter coefficients. The first term represents the current field n and the second term represents the previous field $n-1$. The motion vector $\bar{e}(\bar{x}, n)$ is defined as:

$$\bar{e}(\bar{x}, n) = \begin{pmatrix} d_x(\bar{x}, n) \\ 2\text{Round}\left(\frac{d_y(\bar{x}, n)}{2}\right) \end{pmatrix}$$

with Round () rounding to the nearest integer value and the vertical motion fraction δ_y defined by:

$$\delta_y(\bar{x}, n) = \left| d_y(\bar{x}, n) - 2\text{Round}\left(\frac{d_y(\bar{x}, n)}{2}\right) \right|$$

The GST-filter, composed of the linear GST-filters h_1 and h_2 , depends on the vertical motion fraction $\delta_y(\bar{x}, n)$ and on the sub-pixel interpolator type.

Delonge proposed to just use vertical interpolators and thus use interpolation only in the y-direction. If a progressive image F^p is available, F^e for the even lines could be determined from the luminance values of the odd lines F^o as:

$$F^e(z, n) = (F^p(z, n-1)H(z))_e = F^o(z, n-1)H^o(z) + F^e(z, n-1)H^e(z)$$

in the z -domain where F^e is the even image and F^o is the odd image. Then F^o can be rewritten as:

$$F^o(z, n-1) = \frac{F^o(z, n) - F^e(z, n-1)H^o(z)}{H^e(z)}$$

which results in:

$$F^e(z, n) = H_1(z)F^o(z, n) + H_2(z)F^e(z, n-1)$$

The linear interpolators can be written as:

$$H_1(z) = \frac{H^o(z)}{H^e(z)}$$

When using sinc-waveform interpolators for deriving the filter coefficients, the linear interpolators $H_1(z)$ and $H_2(z)$ may be written in the k-domain:

$$h_1(k) = (-1)^k \sin c\left(\pi\left(k - \frac{1}{2}\right)\right) \frac{\sin(\pi\delta_r)}{\cos(\pi\delta_r)}$$

$$5 \quad h_2(k) = (-1)^k \frac{\sin c\left(\pi\left(k + \delta_r\right)\right)}{\cos(\pi\delta_r)}$$

When using a first-order linear interpolator, a GST-filter has three taps. The interpolator uses two neighboring pixels on the frame grid. The derivation of the filter coefficients is done by shifting the samples from the previous temporal frame to the current temporal frame. As such, the region of linearity for a first-order linear interpolator starts at the position of the motion compensated sample. When centering the region of linearity to the center of the nearest original and motion compensated sample, the resulting GST-filters may have four taps. Thus, the robustness of the GST-filter is increased.

However, current GST-filters do not take into account any pixels situated in the horizontal direction. Only pixels in the vertical vicinity of the samples pixel and from a temporal previous field, e.g. motion compensated, are used for interpolating the pixel samples.

It is therefore an object of the invention, to provide a de-interpolator which is more robust. It is a further object of the invention, to provide a de-interpolator which provides more exact pixel samples.

The inventions solves these objects by providing a method for de-interlacing a video signal, wherein at least a first pixel from said current field of input pixels is weighted depending on a horizontal component of said estimated motion vector for calculating said interpolated output pixel.

The combination of the horizontal interpolation with the GST vertical interpolation in a 2-D inseparable GST-filter results in a more robust interpolator. As video signals are functions of time and two spatial directions, the de-interlacing which treats both spatial directions results in a better interpolation. The image quality is improved. The distribution of pixels used in the interpolation is more compact than in the vertical only interpolation. That means pixels used for interpolation are located spatially closer to the interpolated pixels. The area pixels are recruited from for interpolation may be smaller. The

PHNL031092

PCT/IB2004/051560

4

price-performance ratio of the interpolator is improved by using a GST-based de-interlacing using both horizontally and vertically neighboring pixels.

A motion vector may be derived from motion components of pixels within the video signal. The motion vector represents the direction of motion of pixels within the video image. A current field of input pixels may be a set of pixels, which are temporal currently displayed or received within the video signal. A weighted sum of input pixels may be acquired by weighting the luminance or chrominance values of the input pixels according to interpolation parameters.

Performing interpolation in the horizontal direction may lead, in combination with vertical GST-filter interpolation, to a 10-taps filter. This may be referred to as a 1-D GST, 4-taps interpolator, the four referring to the vertical GST-filter only. The region of linearity, as described above, may be defined for vertical and horizontal interpolation by a 2-D region of linearity. Mathematically, this may be done by finding a reciprocal lattice of the frequency spectrum, which can be formulated with a simple equation:

$$\bar{f}\bar{x} = 1$$

where $\bar{f} = (f_h, f_v)$ is the frequency in the $\bar{x} = (x, y)$ direction. The region of linearity is a square which has the diagonal equal to one pixel size. In the 2-D situation, the position of the lattice may be freely shifted in the horizontal direction. The centers of triangular-wave interpolators may be at the positions $x + p + \delta_x$ in the horizontal direction, with p an arbitrary integer. By shifting the 2-D region of linearity, the aperture of the GST-filter in the horizontal direction may be increased. By shifting the vertical coordinate of the center of the triangular-wave interpolators by $y + m$, an interpolator with 5-taps may be realized. The sampled pixel may be expressed by:

$$\begin{aligned}
 P(x, y, n) = & -\frac{\delta_y |\delta_x| (1 - |\delta_x|) A(x-1, y + \text{sign}(\delta_y), n)}{1 - \delta_y} \\
 & -\frac{\delta_y (|\delta_x|^2 + (1 - |\delta_x|)^2) A(x, y + \text{sign}(\delta_y), n)}{1 - \delta_y} \\
 & -\frac{\delta_y |\delta_x| (1 - |\delta_x|) A(x+1, y + \text{sign}(\delta_y), n)}{1 - \delta_y} \\
 & + \frac{(1 - |\delta_x|) C(x + \delta_x, y + \delta_y, n \pm 1) + |\delta_x| C(x + \delta_x + \text{sign}(\delta_x), y + \delta_y, n \pm 1)}{1 - \delta_y}
 \end{aligned}$$

with A and C being pixels contributing to the sampled pixel.

PHNL031092

PCT/IB2004/051560

5

A method of claim 2 may increase the robustness of the interpolator. Horizontally neighboring pixels may also contribute to the sampled pixel. The interpolation then also depends on horizontally neighboring pixels.

5 A method of claim 3 results in using pixels which are not within the 2-D region of linearity. Thus, the sampled pixel also depends on pixel values which are spatially located apart from the sampled pixel.

According to a method of claim 4, a previous field of input pixels is defined, which means that a temporal previous image is used for defining input pixels. The input pixels of the previous field may be motion compensated by using the motion vector.

10 According to claim 4 the pixel being closest to the sampled pixel when motion compensated is used for calculating the sampled output pixel.

According to claim 5, horizontally neighboring vertical lines may be used for calculating the sampled output pixel. Thus, also a vertical component is used for the sampled output pixel.

15 The sign and the absolute value of the motion vector may be used according to claim 6 and 7.

According to claim 8, where input pixels of a previous field, a next field and a current field are used to calculate first, second and third output pixels and where the final output pixel is calculated based on a weighted sum of these output pixels, temporally and

20 spatially neighboring pixels may be used for calculating the sampled output pixel. This increases the robustness of the de-interlacing.

A method according to claim 9 allows for using a special relationship between input pixels which are temporally separated by a current pixel.

Another aspect of the invention is a display device for displaying a de-

25 interlaced video signal comprising estimation means for estimating a motion vector of pixels, definition means for defining a current field of input pixels from said video signal to be used for calculating an interpolated output pixel, calculation means for calculating an interpolated output pixel from a weighted sum of said input pixels and weighting means for weighting at least a first pixel from said current field of input pixels depending on a horizontal component

30 of said estimated motion vector for calculating said interpolated output pixel.

Another aspect of the invention is a computer program for de-interlacing a video signal operable to cause a processor to estimate a motion vector for pixels from said video signal, define a current field of input pixels from said video signal to be used for calculating an interpolated output pixel, calculate an interpolated output pixel from a

PHNL031092

PCT/IB2004/051560

6

weighted sum of said input pixels, and weight at least a first pixel from said current field of input pixels depending on a horizontal component of said estimated motion vector for calculating said interpolated output pixel.

5

These and other aspects of the invention will be apparent from and elucidated with reference to the embodiments described hereinafter:

Fig. 1 depicts an interpolation according to GST-de-interlacing;

Fig. 2 depicts a first-order linear interpolating;

10

Fig. 3 depicts a region of linearity;

Fig. 4 depicts a position of a region of linearity for an inventive interpolator with horizontal contribution of pixels to the output pixel;

Fig. 5 depicts diagrammatically an inventive method;

Fig. 6 depicts an inventive display device.

15

Fig. 2 depicts the result of a first-order linear interpolator, wherein like numerals as in Fig. 1 depict like elements. As the interpolated sample pixel 10 is a weighted sum of neighboring pixels, the weight of each pixel should be calculated by the interpolator.

20 In case a first-order linear interpolator $H(z) = (1 - \delta_y) + \delta_y z^{-1}$ with $0 \leq \delta_y \leq 1$ the interpolators $H_1(z)$ and $H_2(z)$ may be given as:

$$H_1(z) = \frac{\delta_y}{1 - \delta_y} z^{-1}$$

$$H_2(z) = (1 - \delta_y) - \frac{(\delta_y)^2}{1 - \delta_y} z^{-2}$$

25 The motion vector may be relevant for the weighting of each pixel. In case a motion of 0.5 pixel per field, i.e. $\delta_y = 0.5$, is given, the inverse z-transform of even field

$F^e(z, n)$ results in the spatio-temporal expression for $F^e(y, n)$:

$$F^e(y, n) = F^o(y+1, n) + \frac{1}{2} F^e(y, n-1) - \frac{1}{2} F^e(y+2, n-1)$$

As can be seen from Fig. 2, the neighboring pixels of the previous field $n-1$ are weighted with 0.5 and the neighboring pixel of the current field n is weighted with 1.

30 The first-order linear interpolator as depicted in Fig. 2 results in a three taps GST-filter. The above calculation assumes linearity between two neighboring pixels on the frame grid. In

case the region of linearity is centered to the center of the nearest original and motion compensated sample, the resulting GST-filter may have four taps. The additional tap in these four taps GST-filters increases the contribution of spatially neighboring sample values. Two sets of independent samples from the current field and from previous/next temporal fields, shifted over the motion vector, may be used for GST-filtering only in the vertical direction according the prior art. As the interpolator can only be used on a so-called region of linearity, which has the size of one pixel, the number of taps depends on where the region of linearity is located. This means that up to four neighboring pixels in the vertical direction may be used for interpolation.

As the more pixels are used, the better results are obtained, it should be possible to use more pixels. This may be done by using pixels situated in the horizontal vicinity of the sampled pixel. When using pixels shifted in the horizontal direction, an average value may be used for interpolation, which is:

$$C_{av}(x, y + \delta_y, n \pm 1) = (1 - |\delta_x|)C(x + \delta_x, y + \delta_y, n \pm 1) + |\delta_x|C(x + \text{sign}(\delta_x) + \delta_x, y + \delta_y, n \pm 1)$$

The \pm -sign refers to whether the previous or the next field is used in the interpolation. The combination of such a horizontal interpolation with a vertical GST-filter interpolation allows using a separable 10-taps filter.

To use both pixels in the vertical and horizontal direction, the region of linearity has to be chosen accordingly. In particular in video signals, these are function of time and two spatial directions. Therefore, it is possible to define a de-interlacing algorithm that treats both spatial directions equally.

In case taking horizontally and vertically neighboring pixels into account, the region of linearity may be defined as a grid defining a 2-D region of linearity. This 2-D region of linearity may be found within a reciprocal lattice of the frequency spectrum.

Fig. 3 depicts a reciprocal lattice 12 in the frequency domain and the spatial domain, respectively. The lattice 12 defines the region of linearity which is now a parallelogram. A linear relation is established between pixels separated by a distance $|\bar{x}|$ in the \bar{x} direction. Further, the triangular interpolator used in the 1-dimensional interpolator may take the shape of a pyramidal interpolator. Shifting the region of linearity in the vertical or horizontal direction leads to different numbers of filter taps. In particular, if the pyramidal

PHNL031092

PCT/IB2004/051560

8

interpolators are centered at position $(x + p, y)$, with p an arbitrary integer the 1-D case may result.

In the 2-D situation, the position of the lattice 12 in the horizontal direction may be freely shifted. The simplest shifting may result in centering the pyramids at the position $x + p + \delta_x$ in the horizontal direction, with p an arbitrary integer. This leads to a larger aperture of the GST-filter in the horizontal direction. In case the vertical coordinate of the center of the pyramidal interpolator is $y + m$, a five-taps interpolator may be obtained. The sampled pixel may be expressed by:

$$\begin{aligned}
 P(x, y, n) = & -\frac{\delta_y |\delta_x| (1 - |\delta_x|) A(x-1, y + \text{sign}(\delta_y), n)}{1 - \delta_y} \\
 & -\frac{\delta_y (|\delta_x|^2 + (1 - |\delta_x|)^2) A(x, y + \text{sign}(\delta_y), n)}{1 - \delta_y} \\
 & -\frac{\delta_y |\delta_x| (1 - |\delta_x|) A(x+1, y + \text{sign}(\delta_y), n)}{1 - \delta_y} \\
 & + \frac{C_{av}(x + \delta_x, y + \delta_y, n \pm 1)}{1 - \delta_y}
 \end{aligned}$$

It may be possible, as depicted in Fig. 4, to interpolate pixels which are symmetrically situated to the pixel $P(x, y, n)$. These pixel may be, as depicted in Fig. 4a, $B(x-1, y - \text{sign}(\delta_y), n)$, $B(x, y - \text{sign}(\delta_y), n)$ and $B(x+1, y - \text{sign}(\delta_y), n)$ from the current field. Further from the previous and the next field may be taken $D(x + \delta_x, y - 2\text{sign}(\delta_y) + \delta_y, n \pm 1)$, $D(x + \text{sign}(\delta_x) + \delta_x, y - 2\text{sign}(\delta_x) + \delta_y, n \pm 1)$. As depicted in Fig. 4a, a five-taps interpolator takes into account the above-mentioned pixel values. When shifting the region of linearity in direction of the motion vector, a further value $C(x + \delta_x, y + \delta_y, n \pm 1)$ may be used.

According to the invention, the region of pixels contributing to the interpolation is extended in the horizontal direction. The interpolation results are improved in particular for sequences with a diagonal motion.

Fig. 5 depicts a method according to the invention. In step 50 a motion vector is estimated from an input video signal 48. The input video signal 48 is divided up in regions of linearity in step 52 for a current field, a previous field and a next field. After that, in step 54, horizontally neighboring pixels as well as motion compensated pixels using a horizontal component of the motion vector are weighted according to the motion vector. In step 56, vertically relevant pixels are weighted according to the motion vector.

PHNL031092

PCT/IB2004/051560

In step 58, the weighted pixel values are summed and interpolated, resulting in an interpolated pixel sample. This interpolated pixel sample may be used for creating an odd line of pixels when only even lines of pixels are transmitted within the video signal 48. The image quality may be increased.

5 Fig. 6 depicts a display device 60. An input video signal 48 is fed to said display device 60 and received within a receiver 62. The receiver 62 provides the received images to storage 64. In motion estimator 66, motion vectors are estimated from the video signals. Pixels from the current, the previous and the next field are taken from the storage 64 and weighted in the weighting means 68, in particular according to the estimated motion
10 vector. The weighted pixel values are provided to summer 70, where a weighted sum is calculated. The resulting value is fed to output 72.

 With the inventive method, computer program and display device the image quality may be increased without increasing transmission bandwidth. This is in particular relevant when display devices are able to provide higher resolution than transmission
15 bandwidth is available.